

This is a repository copy of *A Shadow Constrained Conditional Generative Adversarial Net for SRTM Data Restoration*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/154848/>

Version: Accepted Version

Article:

Dong, Guoshuai, Huang, Weimin, Smith, William Alfred Peter orcid.org/0000-0002-6047-0413 et al. (1 more author) (2020) A Shadow Constrained Conditional Generative Adversarial Net for SRTM Data Restoration. *Remote Sensing of Environment*. 111602. ISSN 0034-4257

<https://doi.org/10.1016/j.rse.2019.111602>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

A Shadow Constrained Conditional Generative Adversarial Net for SRTM Data Restoration

Guoshuai Dong^a, Weimin Huang^b, William A. P. Smith^c, Peng Ren^{a,*}

^aCollege of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao 266580, China.

^bFaculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's A1C 5S7, Canada.

^cDepartment of Computer Science, University of York, York YO10 5GH, United Kingdom.

Abstract

The original data produced by the Shuttle Radar Topography Mission (SRTM) tend to have an abundance of voids in mountainous areas where the elevation measurements are missing. In this paper, deep learning models are investigated for restoring SRTM data. To this end, we explore generative adversarial nets, which represent one state of the art family of deep learning models. A conditional generative adversarial network (CGAN) is introduced as the baseline method for filling voids in incomplete SRTM data. The problem regarding shadow violation that possibly arises from the CGAN restored data is investigated. To address this deficiency, shadow geometric constraints based on shadow maps of satellite images are devised. In addition, a shadow constrained conditional generative adversarial network (SCGAN), which incorporates the shadow geometric constraints into the CGAN, is developed. Training the SCGAN model requires both the remote sensing observations (i.e., the original incomplete SRTM data and satellite images) and the ground truth data (i.e., the complete SRTM data, which are manually refined from the incomplete SRTM data with the reference of in-situ measurements). The integration of the multi-source training data enables the SCGAN model to be characterized by comprehensive information including both mountain shape variation and mountain shadow geometry. Experimental results validate the superiority of the SCGAN over the comparison methods, i.e., the interpolation, the convolutional neural network (CNN) and the baseline CGAN, in SRTM data restoration.

Keywords: SRTM data restoration, multi-source data, shadow geometric constraints, shadow constrained conditional generative network.

1. Introduction

In 2000, the National Aeronautics and Space Administration (NASA) and the National Geospatial Intelligence Agency (NGA) operated the Shuttle Radar Topographic Mission (SRTM), which acquired SRTM data (elevation data) covering more than 80% of the earth's surface between the north latitude 60 degree and south latitude 56 degree (Rodriguez et al., 2005, 2006; Farr et al., 2007; Wendleder et al., 2016). The SRTM data

*Corresponding author.

Email address: pengren@upc.edu.cn (Peng Ren)

have been widely used in the field of remote sensing, such as natural disaster monitoring (Jafarzadegan & Merwade, 2017), meteorological forecast (Yue et al., 2015), and glacial evolution detection (Lee et al., 2015; Yang et al., 2015). In addition, digital elevation models (DEMs) were created based on the SRTM data. The highly accurate DEMs are of great significance to topographic mapping (Toutin, 2008), hydrological simulation (Rizzoli et al., 2017), and flood control (Yue et al., 2017). The accuracy of DEM data highly relies on that of the SRTM data. Therefore, there are strict requirements for the SRTM data accuracy. However, numerous voids (missing data) exist in SRTM data, especially in mountainous regions. One reason for this deficiency is that it is difficult for radar to capture steep terrains. Furthermore, the environment in mountainous areas is particularly complex, and there may be numerous environmental sources of noise that interfere with radar signals, resulting in incomplete and blurry data. These missing data regions account for 0.3% of the surveyed area (Kelndorfer et al., 2004; Hall et al., 2005; Hirt, 2018). The completeness and accuracy of the SRTM data are negatively affected by missing data in mountainous areas. One widely used strategy for SRTM data restoration is manually refining the data by consulting alternative observations. Furthermore, the raw SRTM data tend to contain numerous outliers which need to be manually removed. Manually removing outliers may achieve highly accurate terrain restoration. In the condition that the scale of the data is huge and voids are relatively common, the accurate manually manipulated restoration requires intensive human labor. Furthermore, manual refinement is subjective such that the restoration accuracy significantly depends on the person who does the manual work. On the contrary, automatic refinement methods provide an objective means for SRTM data restoration with little human intervention. However, automatic refinement may not outperform the manual refinement with intensive human observations, as long as the automatic strategy is designed in an ad hoc manner with limited representational power. In this scenario, how to automatically restore the SRTM data for mountainous regions in a precise manner is a challenging and important topic in the remote sensing literature. One goal of our work is to develop a novel deep learning based automatic SRTM data restoration strategy, which achieves state of the art performance for automatic SRTM data restoration.

Most existing automatic data restoration methods in the literature repair incomplete SRTM data based on interpolation. In practical void filling tasks, various interpolation techniques have been adopted. Among these, the inverse distance weighting (IDW) interpolation and the Kriging interpolation techniques are widely used for filling voids in general situations especially for those with medium sizes and in relatively flat low-lying areas. Moreover, spline interpolation, triangular irregular network and their extensions are used for addressing the void filling tasks in more specific situations (Reuter et al., 2007). However, interpolation based methods tend to induce biases, resulting in discontinuous and unsmooth boundaries (Grohman et al., 2006). Additionally, terrain characteristics around the incomplete region significantly influence the performance of the interpolation methods. For example, the same interpolation method performs quite differently with

40 respect to different neighboring terrain characteristics. Therefore, it was suggested that interpolation based
 methods should involve the topographical characteristics as cues to repair incomplete terrain (Reuter et al.,
 2007; Heritage et al., 2009; Arun, 2013). In the light of this observation, researchers used various auxiliary
 data to form geological cues for guiding the filling of no-data regions. Ling et al. (Ling et al., 2007) were
 among the first to interpolate void data according to valley trends in satellite images. However, it is not
 45 clear how this valley interpolation method can be extended to mountainous situations. Subsequently, quite
 a few interpolation methods involving auxiliary DEM data (Reuter et al., 2007; Karkee et al., 2008; Milan
 et al., 2011) were proposed. However, these methods are not scalable to DEM data with different resolutions
 and the auxiliary DEM data are not always available. Different from using DEM as auxiliary data, Hogan et
 al. (Hogan & Smith, 2010) proposed to refine interpolation results by using geometric shadow constraints.
 50 However, shadow cues alone provide a relatively sparse source of information, and the method involves solving
 a difficult and non-convex optimization problem which is computationally expensive.

Though the information extracted from auxiliary data improves the interpolation accuracy, it only fo-
 cuses on local data characteristics. Generally, most existing interpolation based methods neither consider
 the global information contained in the non-void regions nor are they capable of learning heuristics from
 55 the global data variation to fill the voids. These limitations can possibly be addressed by exploiting deep
 learning models. Most deep learning models (e.g., convolutional neural networks (CNNs)) require a large
 amount of training data. Deep models learn extensively from big data and produce effective representations
 characterized by global and local features that have been exhibited. Therefore, deep learning models are
 capable of characterizing the global information based on non-void data and deriving terrain variation trends
 60 for void filling.

In contrast to most traditional deep learning methods such as CNNs that only model one unique net, a new
 family of deep learning models, i.e., generative adversarial nets (GANs) (Goodfellow et al., 2014), has recently
 been proposed in the literature. Specifically, one GAN model consists of two nets, i.e., a generator and a
 discriminator. The generator takes noise as inputs and tries its best to generate data (e.g., images) as realistic
 65 as ground truth data. The discriminator, on the other hand, tries its best to discriminate the generated data
 from the ground truth. The two nets play an adversarial game and their capabilities of generating realistic
 data and discriminating generated data are both maximized during the adversarial training process. The
 adversarial training is terminated when the generated data appear so realistic that the discriminator cannot
 distinguish them from the ground truth. Compared with traditional deep learning models such as CNNs, a
 70 GAN has two advantages which make it a promising means for SRTM restoration. One is that a GAN does
 not necessarily require a large amount of training data, because in the training procedure it generates new
 data that enlarge the training dataset. The other advantage is that a GAN not only memorizes information
 extracted from training data but also learns to generate novel but plausible new instances that appear to be

drown from the same distribution as the real data. These two advantages reveal the potential of GANs as
75 a means for SRTM data restoration. It is common in practice that only a small set of the void and ground
truth patch pairs are available and GANs can be trained with a small size training dataset. Meanwhile, unlike
CNNs which restore data by retrieval from memorized information, GANs have the capability of imagination
and can derive heuristics from training data and restore terrain data in a more comprehensive manner.

To take advantage of the representational power of GANs, an SRTM restoration framework is developed
80 based on GANs. Rather than employing its original form which takes noise as the generator’s inputs, we
commence by investigating the possible way for SRTM data restoration based on conditional generative
adversarial nets (CGANs). Specifically, regions containing voids are taken as the inputs of the generator of
a CGAN and such conditional data reduce the ambiguities arising from the noise inputs. In addition, the
complete SRTM data are taken as the target outputs of the generator of the CGAN and thus the generator
85 and discriminator are adversarially trained in a supervised fashion. The void filling CGAN was previously
investigated in our preliminary work (Dong et al., 2018). It is considered as one baseline method in this
paper.

However, the baseline CGAN has limitations for void filling. First, it is a straightforward data-driven
approach without considering the geometric characteristics of terrains. It may restore the terrain in a way
90 that just follows terrain variation heuristics but violates certain terrain geometric constraints. Second, the
baseline CGAN is dependent on the quantity and quality of complete SRTM data that can be provided.
The training samples themselves may contain errors or may be produced by a separate void filling procedure
and so may also lead to restorations that violate geometric constraints. Therefore, the baseline CGAN alone
cannot guarantee that the restored data follow the geometric nature of terrains. This deficiency motivates
95 us to develop geometric cues for constraining the CGAN for the purpose of accurate SRTM data restoration
subject to terrain topological nature. One such geometric property arises from observing cast shadows in
terrain imageries. We make an attempt to introduce additional cast shadow supervision to the CGAN.
This shadow constrained conditional generative adversarial net (SCGAN) encourages the restored SRTM
data to adhere to shadow geometry constraints, potentially leading to improvements in the non-void regions.
100 Compared with existing interpolation based methods, our SCGAN model learns a powerful representation
of SRTM data and comprehensively characterizes the relationship between missing data and valid data. It
excludes the need of auxiliary DEMs and avoids the problem of ambiguous resolutions from different DEMs.
Furthermore, in contrast to most existing deep learning models such as CNNs and the baseline CGAN which
are completely driven by data, our SCGAN takes into account the effects of mountain shadow conditions
105 and accordingly constructs geometric constraints based on shadow maps to guide the training process of the
SCGAN model. It thus has the potential of avoiding the inaccurate restoration that violates terrain shadow
geometry.

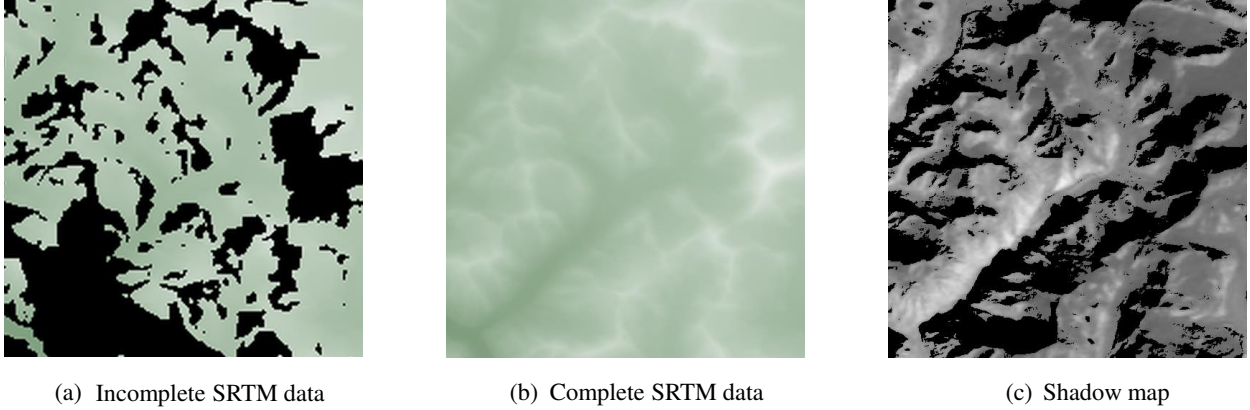


Fig. 1. A sample of data.

The organization of the paper is as follows. Section 2 describes the data catalog of our study areas. Section 3 introduces the baseline CGAN model for SRTM data restoration and discusses its limitations. Section 4 investigates the shadow geometrics for mountainous regions and Section 5 develops the shadow constraints with respect to the shadow geometrics. Section 6 describes how to incorporate the shadow constraints as supervision information and presents the shadow constrained conditional generative adversarial net (SCGAN). Section 7 describes the restoration procedure based on the SCGAN. Section 8 presents and analyzes the empirical evaluations of the SCGAN and comparison methods. Section 9 concludes our work.

2. Data catalog for study areas

Training our proposed SCGAN model for SRTM data restoration requires three types of data for the same mountainous regions, i.e., (a) incomplete SRTM data, (b) complete SRTM data and (c) shadow maps. A region in the southwest of China is selected as the study area. The region covers a geographical rectangle from the northwest corner $29^{\circ}N85^{\circ}E$ to the southeast corner $28^{\circ}N86^{\circ}E$. The required data are obtained for the region.

There are multiple versions of SRTM data available. The original SRTM data (SRTM version 1) have a lot of voids but lack the indications of peaks, water bodies and coastlines. Therefore, the data are not appropriate for scientific research and remote sensing missions. The data of SRTM version 2¹ indicate peaks, water bodies and coastlines (Boncori, 2016), but still have plenty of voids. Most of the missing data of SRTM version 3 were filled except for some small regions in the mountainous areas. SRTM version 4.1 completed by the International Center for Tropical Agriculture (CIAT) produces the seamless SRTM data without data missing (Jarvis et al., 2008).

Taking into account the qualities of different versions of SRTM data, the data from SRTM version 2 are taken as (a) incomplete SRTM data, and those from SRTM version 4.1 are taken as (b) complete SRTM

¹https://dds.cr.usgs.gov/srtm/version2_1/

130 data. In addition, (c) shadow maps covering the same regions as SRTM data are extracted from satellite images. Fig. 1 displays sample patches of (a) an incomplete SRTM data patch, (b) a complete SRTM data patch, and (c) a shadow map patch for the same region.

The three types of data (a), (b) and (c) are obtained through different means. The original incomplete SRTM data, i.e., the data (a), are obtained from shuttle shipped radar. The data from SRTM version 4.1, i.e., 135 the data (b), are refined from the original incomplete SRTM data with the reference of in-situ measurements. The shadow maps, i.e., the data (c), are obtained by processing the Landsat-5 satellite multispectral images. Training our proposed SCGAN model involves both aerial and in-situ measurements. Therefore, the SCGAN not only provides a new deep learning model that incorporates shadow geometrics but also forms a novel data fusion framework that integrates satellite, shuttle based radar and in-situ observations data.

140 Furthermore, one alternative way for looking at the SRTM data types is from the perspective of spatial resolution. In this scenario, two widely used SRTM data types are SRTM 3 Arc-Second global data with the resolution of 90 m and the SRTM 1 Arc-Second global data with the resolution of 30 m, separately. It is worth noting that one mission goal of SRTM is to attain an altitude accuracy in terms of root mean square error (RMSE) within the order of magnitude 10 m (Mukul et al., 2017). Though the SRTM 1 Arc-Second 145 dataset provides data with acceptable altitude accuracy for most of the relatively flat low-lying areas, the goal for accurately completing high altitude voids is still challenging. Most existing interpolation methods cannot achieve the accuracy specified in the SRTM mission goal for restoring missing high altitude data.

In our work, we will forward the state of the art research one step further by developing the deep learning based SCGAN restoration strategy. Experimental evaluations show that our new method is capable of 150 obtaining a restoration accuracy closer to the SRTM mission goal based on different types of data used here. More specifically, experimental results validate that our new strategy improves not only the performance for enhancing the quality of the high resolution SRTM 1 Arc-Second data but also the interpolation accuracy for the low resolution SRTM 3 Arc-Second dataset. Therefore, the SCGAN is a general methodology for remote sensing surface restoration. It is not confined to filling voids for one specific type of remote sensing data but 155 can be generalized to restoring various remote sensing data.

3. A Baseline Conditional Generative Adversarial Net for Filling SRTM Voids in Mountainous No-data Areas

This section introduces our baseline conditional generative adversarial net (CGAN) for SRTM mountainous data restoration. We commence by presenting the basic structure of the CGAN and describe how to apply it 160 to the task of filling SRTM voids in mountainous no-data areas. The limitations of the baseline CGAN are discussed at the end of the section.

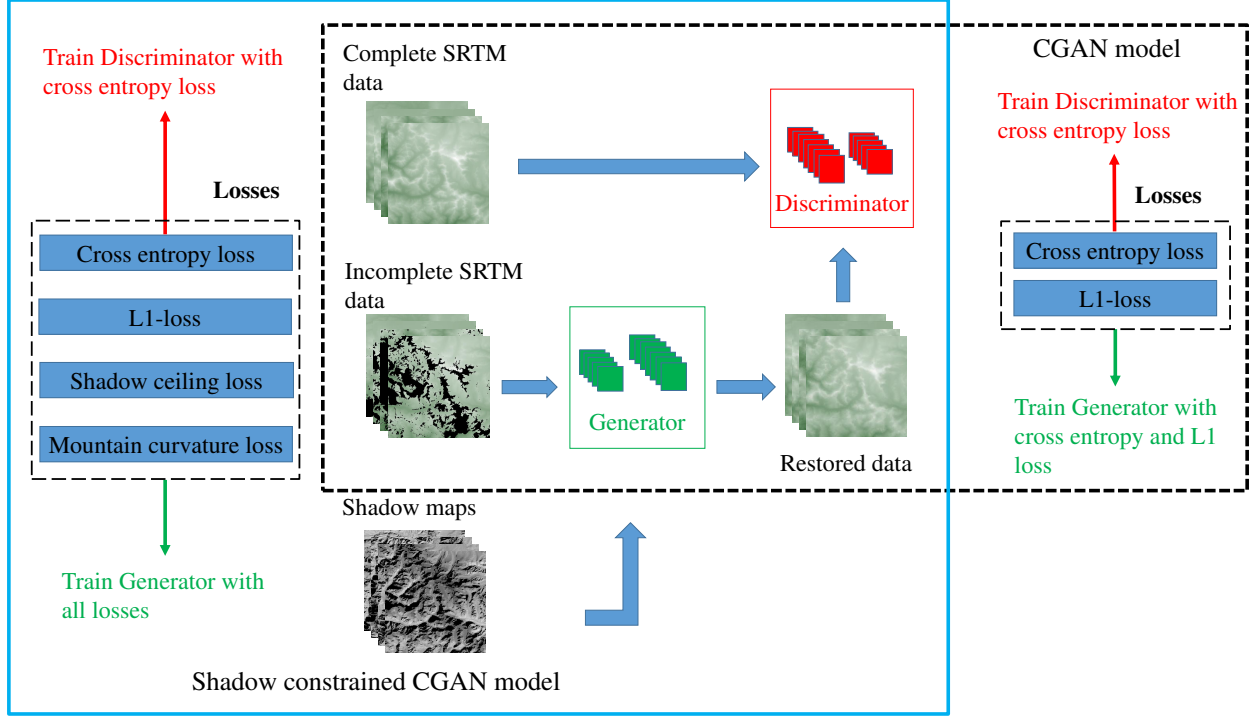


Fig. 2. SRTM data restoration by the CGAN and SCGAN models.

3.1. Conditional Generative Adversarial Nets

Generative Adversarial Networks (GANs) (Goodfellow et al., 2014; Creswell et al., 2018; Zhan et al., 2018) have shown great effectiveness in learning the distribution of a dataset and generating new data. However, the basic GAN takes random noise as the generator’s inputs which lead to ambiguity in generating new data. Additionally, the original GAN also suffers from locally optimal solution and gradient disappearance that possibly lead to considerable inaccuracy. Our motivation is to train the generator of a GAN to generate data for void regions. In this scenario, one input to the generator is a data patch consisting of both valid and void regions, and the generator is expected to fill the void regions and output a restored complete data patch. Therefore, in contrast to the basic GAN, the generator is trained with incomplete data as inputs rather than random noise. This further motivates us to exploit a more comprehensive GAN structure, i.e., conditional generative adversarial net (CGAN) (Isola et al., 2017; Antipov et al., 2017), for the purpose of restoring the SRTM data in no-data mountainous areas. Specifically, the valid regions in the incomplete data patches provide conditioning for guiding the data generation process. The CGAN model takes such conditioning input as latent codes which address the lack of control in the basic GAN. The CGAN structure is introduced in the next subsection.

3.2. Filling SRTM Voids via Conditional Generative Adversarial Nets

The conditional generative adversarial net (CGAN) for SRTM data restoration is illustrated within the black dash rectangle in Fig. 2. The CGAN model comprises a generator G and a discriminator D . Both the

180 generator G and the discriminator D are structured in terms of the U-net deep learning structure (Gatys et al., 2016; Shelhamer et al., 2017; Yu et al., 2018), separately.

The generator G takes incomplete SRTM data (with elevation in void regions set to be zero) as inputs and outputs the restored SRTM data. An incomplete SRTM data patch and its restored SRTM data patch produced by the generator are represented as the matrices H_V and H_G , respectively. The input-output
185 relation of the generator G is formulated as follows:

$$H_G = G(H_V). \quad (1)$$

The ground truth SRTM data patch is represented as the matrix H , which covers the same region as that of H_V . The discriminator D takes both a restored data patch H_G and its ground truth data patch H as an input and outputs a score $D(H_G, H)$ that determines the dissimilarity between H_G and H in terms of a score value in the interval $[0,1]$.

190 The generator G aims at producing the restored SRTM data H_G so as to resemble the ground truth data H as best as possible. This is effected by training G with respect to minimizing two loss functions, which evaluate the similarity from two perspectives separately. The first loss is an ℓ_1 norm between H_G and H :

$$L_1 = |H_G - H|. \quad (2)$$

L_1 provides an objective similarity measure between H_G and H . The second loss provides a similarity measure between H_G and H obtained from the discriminator as follows:

$$L_G = E[\log(1 - D(H_G, H))]. \quad (3)$$

195 As $D(H_G, H)$ indicates the dissimilarity between H_G and H , the term $1 - D(H_G, H)$ represents the similarity between them judged by D . The expectation and logarithm operations are taken to form a principled formulation. In this scenario, the loss L_G measures the similarity between H_G and H from the view of discriminator D . Minimizing L_G encourages the generator G to fool the discriminator D as much as possible.

200 In addition, the goal of the discriminator D is to try its best to discriminate generated data from ground truth data. To this end, the discriminator D is trained by maximizing the following loss:

$$L_D = E[\log D(H_G, H)]. \quad (4)$$

The loss L_D is the expectation of the logarithm of the dissimilarity score provided by D . Training D by maximizing the loss L_D increases the capability of D for discriminating data generated by G from ground truth data.

205 The sum of the losses (2) and (3) for training the generator G and the loss (4) for training the discriminator

D forms the overall loss for training the CGAN as follows:

$$L_{\text{CGAN}} = \alpha L_1 + L_G + L_D, \quad (5)$$

where α is an empirical scaling parameter. The expectation and logarithm operations enable (3) to be easily incorporated into the principled formulation in terms of a cross entropy loss as follows:

$$L_D + L_G = E[\log D(H_G, H)] + E[\log(1 - D(H_G, H))]. \quad (6)$$

Therefore, the loss L_{CGAN} can be considered as the combination of a cross entropy loss and an ℓ_1 loss.

210 Training G and D leads to the minimization and maximization of the loss L_{CGAN} , respectively. Therefore, the optimal G_* and D_* are obtained by the min-max optimization of the loss L_{CGAN} in (5) as follows:

$$(G_*, D_*) = \arg \min_G \max_D L_{\text{CGAN}}. \quad (7)$$

Algorithm 1: Training procedure for the CGAN model.

Input: Incomplete SRTM data H_V , complete SRTM data H .

for the number of training iterations **do**

Update G :

 Take mini-batch examples from input to do restorations;

 Update parameters of G via minimizing the sum of (2) and (3) by using the Adam optimizer.

Update D :

 Take restored data H_G of G and complete data H as inputs;

 Update parameters of D via maximizing (4) by using the Adam optimizer.

end

Output: G_* and D_* .

The min-max training procedure with respect to (7) is an adversarial training process, which can be thought as a competition game between the generator G and the discriminator D . The two nets thus play the adversarial game and their capabilities of generating realistic data and discriminating generated data are both maximized. Although the discriminator D tries its best to distinguish H_G from H , it is the generator G which finally wins the game, resulting in an equilibrium in which the discriminator D is unable to differentiate between the restored data and the ground truth data. Therefore, the generator G_* optimized by the adversarial training process is believed to have the capability of generating realistic SRTM data for filling the data voids. The training process uses the Adam optimizer (Zhu et al., 2017) for computing the minimization and maximization. The overall training process for the CGAN is shown in Algorithm 1. The CGAN is employed as a baseline for restoring realistic SRTM data.

3.3. Limitations

The baseline CGAN model has great potential in restoring SRTM data because of its effective representational power. Our preliminary study (Dong et al., 2018) has validated its effectiveness over the interpolation based methods. However, the baseline CGAN has limitations in void filling. Firstly, it is a straightforward data driven approach without the consideration of terrain geometric characteristics. It restores terrains in a way that follows terrain variation heuristics but may violate terrain geometric constraints. For example, as illustrated in Fig. 3, it is very likely for the CGAN to restore the void region, which is in fact concave, into an incorrect convex shape. This is because the CGAN trained based on valid and ground truth data is prone to infer an unknown shape from valid adjacent data. For the case in Fig. 3, it is more likely for the CGAN to fill the void with a convex shape because the void region is adjacent to a convex peak. However, this restoration obviously violates the true terrain geometry. Secondly, the performance of the baseline CGAN depends on the quantity and quality of complete SRTM data that are available. The training samples themselves may contain errors or may also lead to restorations that violate the terrain geometry if they are produced by a separate void filling procedure. Therefore, the baseline CGAN alone cannot guarantee that the restored data follow the geometric nature of terrains.

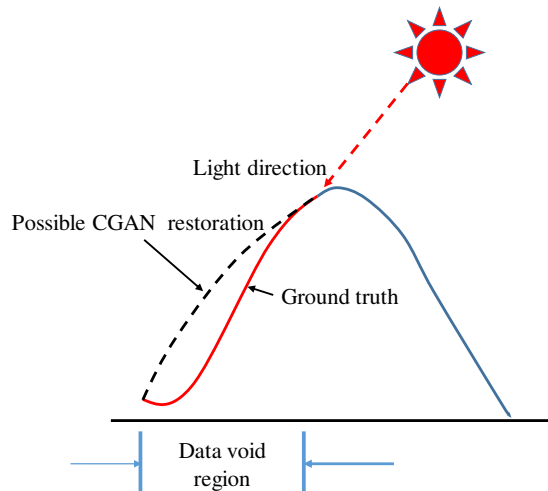


Fig. 3. Limitations of the CGAN model. The incorrect restoration would be apparent if cast shadows were not or inappropriately considered.

The cast shadows provide useful information for improving the CGAN restoration accuracy. For example, in Fig. 3, according to the light direction, the void region is in a shadow. If the void region is incorrectly restored into a convex shape (the black dash curve in Fig. 3), a large part of or even the whole restored shape would be illuminated and this violates the shadow cues. In this case, restoration with a concave shape resembling the ground truth (the red solid curve in Fig. 3) is supposed to satisfy the shadow condition. Furthermore, the shadow cues also provide a remedy for correcting the errors induced by the low-quality

SRTM data. Therefore, in the following sections, we will investigate the shadow geometrics and incorporate
 245 the shadow cues into the CGAN SRTM data restoration model.

4. Shadow Geometry

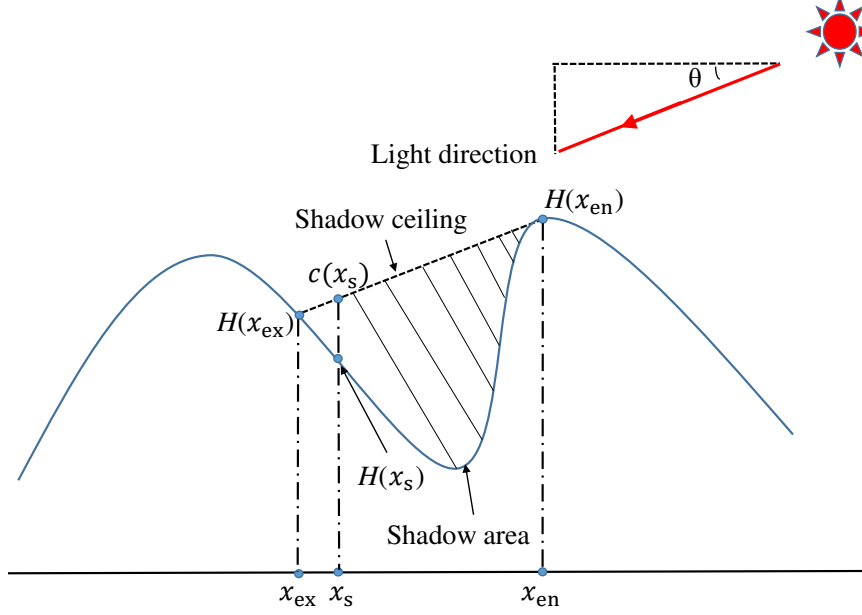


Fig. 4. Mountain shadow geometry.

Any region with large altitude variations (e.g., mountainous areas) contains locations where the sun is
 occluded when not directly overhead. These cast shadow regions provide informative geometric cues that
 can aid SRTM void data filling. The basic geometry of shadowing is illustrated in Fig. 4. The sun is treated
 250 as a point source. Let $\mathbf{s} = [s_1, s_2, s_3]$ be a unit length vector indicating the sun light direction. The angle
 between the light direction and the ground plane is denoted by $\theta = \arccos(s_3)$, where s_3 is the third entry of
 \mathbf{s} and represents the vertical component of the normalized sun light direction vector \mathbf{s} . Consider a 2D slice
 through the DEM that is parallel to both the light direction and the up vector (as illustrated in Fig. 4). A
 pair of locations $(x_{\text{en}}, x_{\text{ex}}) \in X_{\text{sbound}} \subset X \times X$ are defined as a shadow entrance point and a shadow exit
 255 point, respectively. Here $X \times X$ contains all location pairs and X_{sbound} contains all pairs of shadow boundary
 locations.

The shadow ceiling is the line connecting the shadow entrance and exit points. Let $H(x)$ denote the
 altitude of terrain at the location x . It is clear that the shadow ceiling and light direction are parallel, which
 satisfies:

$$\frac{H(x_{\text{en}}) - H(x_{\text{ex}})}{\|x_{\text{en}} - x_{\text{ex}}\|} = \tan(\theta). \quad (8)$$

260 The relationship between the elevations of the shadow boundary and the light direction in (8) leads to

the first shadow cue (C1) in our work:

$$\forall(x_{\text{en}}, x_{\text{ex}}) \in X_{\text{sbound}}, \quad \frac{H(x_{\text{en}}) - H(x_{\text{ex}})}{\|x_{\text{en}} - x_{\text{ex}}\|} - \tan(\theta) = 0. \quad (9)$$

Furthermore, the region below the shadow ceiling lies in a cast shadow and this shadow area is denoted as the set of locations X_s . The elevation of the shadow ceiling at a location $x_s \in X_s$ in the shadow area is represented as $c(x_s, H)$ which is given by:

$$c(x_s, H) = \frac{H(x_{\text{ex}})\|x_s - x_{\text{en}}\| + H(x_{\text{en}})\|x_s - x_{\text{ex}}\|}{\|x_s - x_{\text{en}}\| + \|x_s - x_{\text{ex}}\|}. \quad (10)$$

265 The altitude $H(x_s)$ at any location x_s within a shadow area must be lower than the ceiling elevation $c(x_s, H)$. This observation provides the second shadow constraint (C2) in our work:

$$\forall x_s \in X_s, \quad H(x_s) < c(x_s, H). \quad (11)$$

In addition, the terrain must be convex along the light source direction at a shadow entrance point. Let $\bar{s} = [\frac{s_1}{\sqrt{s_1^2 + s_2^2}}, \frac{s_2}{\sqrt{s_1^2 + s_2^2}}]^T$ denote the orthogonal projection of \mathbf{s} onto the ground plane. The second directional derivative of H along direction \bar{s} must be negative. Using a finite difference approximation of the second
270 derivative, we obtain:

$$H''_{\bar{s}}(x) \approx H(x + \bar{s}) + H(x - \bar{s}) - 2H(x). \quad (12)$$

The convexity constraint results in the third shadow cue (C3) in our work:

$$\forall(x_{\text{en}}, x_{\text{ex}}) \in X_{\text{sbound}}, \quad H''_{\bar{s}}(x_{\text{en}}) < 0. \quad (13)$$

5. Shadow Constraints

We show how to reframe the shadow cues obtained from the previous section as differentiable loss functions for the use within a machine learning scheme.

275 5.1. Shadow segmentation

Shadow segmentation of an image is normally conducted based on pixel intensities under the light source (Chandraker et al., 2007; Wu et al., 2010; Liao et al., 2017; Villarini et al., 2017). The real shadow areas are supposed to be represented by the zero-intensity pixels in an image. However, affected by inter-reflections and ambient illumination, shadow areas are often represented by low (non-zero) pixel intensities. Therefore,
280 one way for detecting shadow areas is to compare the pixel intensity one by one with a global threshold (Lu & Drew, 2005; Chung et al., 2009). Additionally, different from pixel intensity based shadow segmentation methods, some researchers constructed automatic shadow segmentation models with respect to the physical phenomenon of sunlight scattering (Polidorio et al., 2003). Furthermore, in recent years, multispectral images

have been widely used for shadow segmentation in the research field of remote sensing. Shadow segmentation methods according to the intrinsic association between multi-band data and shadow regions (Hogan & Smith, 2010; Teke et al., 2011) have also been developed. Multispectral images characterize shadow information in a more comprehensive manner than single channel densities and shadow segmentation methods based on multispectral images are considered as state of the art methods.

Taking into account the accuracy and robustness of the shadow segmentation methods reviewed above, the multi-band thresholding technique is employed (Hogan & Smith, 2010) to perform shadow segmentation. Specifically, three bands (near infrared, mid-infrared and thermal infrared bands) of multispectral satellite images from Landsat-5 are used for shadow segmentation. Let $I_k(x)$ denote the normalized pixel intensity at x in the k -th band. The segmentation indicator $F(x)$ is formulated as follows:

$$F(x) = \prod_{k=1}^K (1 - I_k(x))^{\varphi_k}, \quad (14)$$

where φ_k is an empirical parameter. A shadow threshold η is applied to the segmentation indicator $F(x)$, resulting in the location set X_s of the shadow region as follows:

$$x = \begin{cases} \in X_s, & \text{if } F(x) > \eta; \\ \notin X_s, & \text{otherwise.} \end{cases} \quad (15)$$

Conducting the thresholding operation (15) for every pixel results in a shadow map of the same size as the image, as shown in Fig. 5. In our work, the black regions in a shadow map represent shadow regions, and the non-black regions exhibit terrain information.

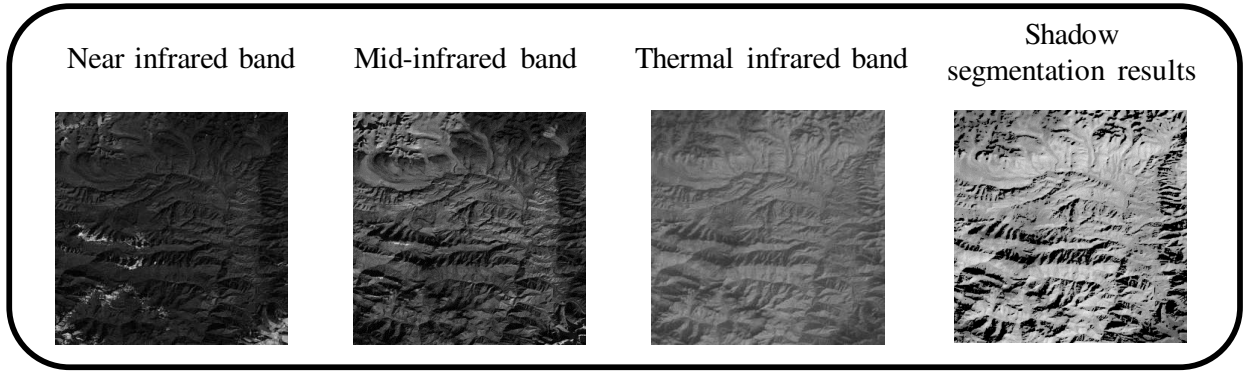


Fig. 5. Shadow segmentation.

5.2. Shadow geometric constraints for SRTM data restoration

Based on the shadow cues introduced in Section 4 and the shadow maps obtained in Section 5.1, we develop shadow geometric constraints for SRTM data restoration. In order to integrate shadow constraints into a machine learning scheme, the shadow constraints are represented in the form of loss functions. Specifically, we

design functions that take on a large positive value when a shadow constraint is violated and zero otherwise.

$H_G(x)$ is denoted as the restored elevation at a geographical point with the location x . The first shadow
 305 cue (C1) in (9) is that the altitudes of shadow boundary satisfy (8). Therefore, the shadow boundary loss
 function L_B is proposed to penalize the situation that X_{shadow} of H_G violates the shadow cue (C1) as follows:

$$L_B = \left| \frac{|H_G(x_{\text{en}}) - H_G(x_{\text{ex}})|}{||x_{\text{en}} - x_{\text{ex}}||} - \tan(\theta) \right|. \quad (16)$$

Furthermore, the second shadow cue (C2) in (11) leads to a positive loss in the case that a restored
 altitude in a shadow area is higher than the corresponding shadow ceiling elevation. An indicator function
 is used to characterize the shadow cue (C2) as follows:

$$H_\varepsilon[H_G(x_s) - c(x_s, H)] = \frac{1}{2} \left[1 + \frac{2}{\pi} \arctan\left(\frac{H_G(x_s) - c(x_s, H)}{\eta}\right) \right]. \quad (17)$$

310 where the parameter η is empirically set to be one in most situations (Chan et al., 2000; Li et al., 2008). This
 results in a value of one if the restoration violates the shadow cue (C2) and zero otherwise. The indicator
 function (17) is used to enhance the disagreement penalty between the restored SRTM data and the true
 SRTM data, and obtain the shadow ceiling loss function L_c with respect to the shadow cue (C2) as follows:

$$L_C = \sum_{x_s \in X_s} H_\varepsilon(H_G(x_s) - c(x_s, H)). \quad (18)$$

Additionally, according to the shadow cue (C3) in (13), a large loss is imposed if the restored shadow
 315 entrance point is located at a valley rather than a peak. Following (12), a convexity characterization function
 is defined as follows:

$$t(H_G, H) = \frac{H_G(x_{\text{en}} + 1) + H_G(x_{\text{en}} - 1)}{2} - H(x_{\text{en}}). \quad (19)$$

Accordingly, another indicator function is denoted as follows:

$$H_\varepsilon(t(H_G, H)) = \frac{1}{2} \left[1 + \frac{2}{\pi} \arctan\left(\frac{t(H_G, H)}{\mu}\right) \right], \quad (20)$$

where the parameter μ is empirically set to be one in most situations (Chan et al., 2000; Li et al., 2008). This
 results in a value of one if the restoration violates (13) and zero otherwise. The shadow entrance curvature
 320 loss function L_v with respect to the shadow cue (C3) is formulated as follows:

$$L_V = \sum_{x_{\text{en}} \in X_{\text{en}}} H_\varepsilon(t(H_G, H)). \quad (21)$$

6. The Shadow Constrained Conditional Generative Adversarial Net

The baseline CGAN model is able to restore SRTM data with the great capability of feature representation and generation. However, as discussed in Section 3.3, the CGAN model does not take into account terrain geometry and may generate irrational restoration that violates shadowing conditions. To address this problem, geometric shadow constraints, i.e., the loss functions (16), (18) and (21), are incorporated into the CGAN loss to achieve shadow guided training. This strategy is referred to as the shadow constrained conditional generative adversarial neural network (SCGAN). Specifically, the overall loss function for the SCGAN is denoted as follows:

$$L_{SCGAN} = L_{CGAN} + \lambda_B L_B + \lambda_C L_C + \lambda_V L_V, \quad (22)$$

where the parameters λ_B , λ_C , and λ_V balance the effects of different terms in the overall loss function. The training process of the SCGAN model is an adversarially (min-max) game between the generator G and the discriminator D . Therefore, the final optimal \hat{G} and \hat{D} are obtained by the min-max game as follows:

$$(\hat{G}, \hat{D}) = \arg \min_G \max_D L_{SCGAN}. \quad (23)$$

The conditions for our SCGAN arise from two folds. Firstly, different from the original GAN whose generator is fed with noise, our generator takes in an incomplete patch with voids and produces a complete patch. Despite the voids, the input patch has non-void data which provides a guiding condition of terrain adjacency for the generator to fill the voids. Secondly, we equip the SCGAN with terrain shadow cues, which enable shadow geometric conditions for the adversarial learning of the SCGAN model. The structure of the SCGAN is illustrated inside the blue solid line box in Fig. 2. The SCGAN model takes incomplete SRTM data as inputs and complete SRTM data as targets with shadow maps as additional supervision. It employs the same U-net deep structures as the original CGAN. The SCGAN model learns the data variation in SRTM data with respect to losses enhanced by geometric shadow constraints, which enable the SCGAN to encode knowledge of shadow cues. Therefore, unlike the baseline CGAN which is solely driven by training data, our SCGAN not only learns from valid data but also follows geometric rules. Especially, the SCGAN has the potential of avoiding the incorrect CGAN data restoration that violates the shadow cues. It thus potentially results in more effective restoration than the baseline CGAN.

7. SRTM Data Restoration via The Shadow Constrained Conditional Generative Adversarial Net

7.1. Training The SCGAN

A training data sample comprises a complete SRTM data patch, an incomplete SRTM data patch and a shadow map for the same region. Fig. 1 shows one training data sample. Training the SCGAN requires

a dataset consisting of the training data samples. The training process is a min-max game between the generator G and the discriminator D . The generator G takes incomplete SRTM data H_V as inputs and generates restored SRTM data H_G . To make the restored SRTM data H_G more realistic, the generator G seeks to minimize loss functions (2) and (3). Furthermore, computing the optimal generator \hat{G} involves the supervision in terms of minimizing loss functions (16), (18) and (21) in addition to (2) and (3). The additional supervision makes the restored SRTM data H_G satisfy the shadow cues (C1), (C2) and (C3). Additionally, the discriminator D tries its best to distinguish the differences between the restored SRTM data H_G and the complete SRTM data H by maximizing loss functions (4). The training process employs the Adam optimizer (Zhu et al., 2017) for computing the minimization and maximization. The Adam optimizer makes the training of the SCGAN model converge fast with economical computing resources. In addition, the Adam optimizer employs an independent adaptive learning rate strategy, which enables the computation for large-scale parameter optimization even more efficient. The training procedure is shown in Algorithm 2.

Algorithm 2: Training procedure for the SCGAN model.

Input: Incomplete SRTM data H_V , complete SRTM data H and shadow maps.

for the number of training iterations **do**

Update G :

 Take mini-batch examples from input to do restorations;

 Update parameters of G via minimizing the sum of (2), (3), (16), (18) and (21) by using the Adam optimizer.

Update D :

 Take restored data H_G of G and complete data H as inputs;

 Update parameters of D via maximizing (4) by using the Adam optimizer.

end

Output: \hat{G} and \hat{D} .

When the discriminator D cannot distinguish the differences between H_G and H , the generator G wins the min-max game. In this scenario, the well trained generator \hat{G} is believed to have the capability of restoring accurate SRTM data from incomplete SRTM data.

7.2. SRTM Data Restoration Based on The SCGAN

The trained generator of the SCGAN takes incomplete SRTM data as inputs and results in the restored SRTM data. The restoration diagram is shown in Fig. 6. Complete SRTM data and shadow maps are not required in the restoration procedure. The information conveyed by complete SRTM data and shadow maps plays an important supervision role in the training process. It encourages the SCGAN to restore the SRTM data as close to the complete SRTM data subject to the shadow constraints as possible. The proposed SCGAN is a parametric model. The trained generator of the SCGAN restores SRTM data without the requirement of either complete SRTM data or shadow maps. Benefiting from the supervised training with respect to the

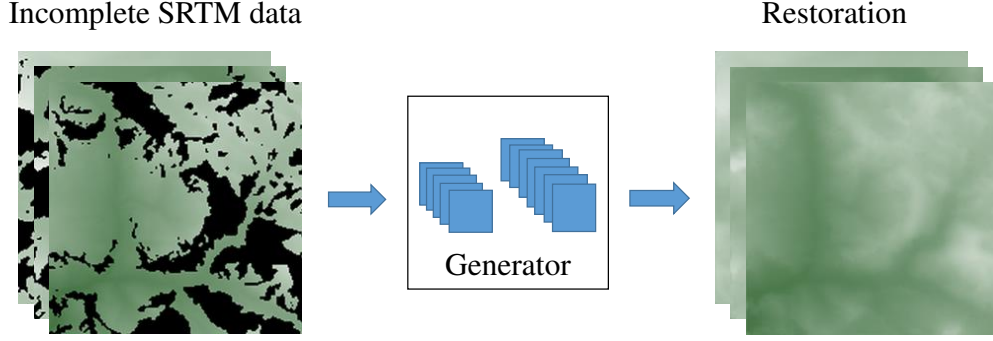


Fig. 6. SRTM data restoration via the generator of the SCGAN.

complete data and shadow maps, the restoration is encouraged to both resemble the ground truth and satisfy the shadow constraints.

8. Experimental evaluations

8.1. Evaluations based on SRTM 3 Arc-Second global data

The resolution of the SRTM 3 Arc-Second global data used in the experiments is 90 m. It is observed that there is a correlation between elevation error and altitude in the SRTM 3 Arc-Second global data for our investigated mountainous areas (Zhang et al., 2018). Specifically, for the region with an altitude less than 1000 m, the elevation error is generally between 10 m and 20 m. For the region with an altitude around 4000 m, the elevation error is more than 40 m. For complex terrains with even higher altitudes, the elevation error can be larger than 100 m. Additionally, noisy radar signals may lead to a standard deviation of SRTM data error greater than 100 m in some mountainous areas. The altitudes of the investigated areas are higher than 4,000 m, and the SRTM 3 Arc-Second global data involved in training or testing may suffer an error larger than 100 m. In the experiments, we use the widely used Kriging interpolation and IDW interpolation as representative interpolation methods for comparison. We empirically compare the void filling results obtained from the Kriging interpolation, the IDW interpolation, the CNN, the baseline CGAN and the proposed SCGAN. Specifically, thirty six non-overlapping scenes are used to evaluate different methods. For the three learning based models, i.e., the CNN, the baseline CGAN and the proposed SCGAN, cross validations are performed by using thirty three scenes and three scenes for training and testing, respectively. We use manually corrected SRTM data (SRTM version 4.1) and shadow maps as labelled training data to train the three learning based models. Then we use incomplete SRTM data as testing data to empirically evaluate the performance of the three learning based models for data restoration. It should be noted that the training data and the testing data are completely different data covering non-overlapped regions. Additionally, the three learning based models do not see any testing data during the training process, and the training data are not used for testing the models.

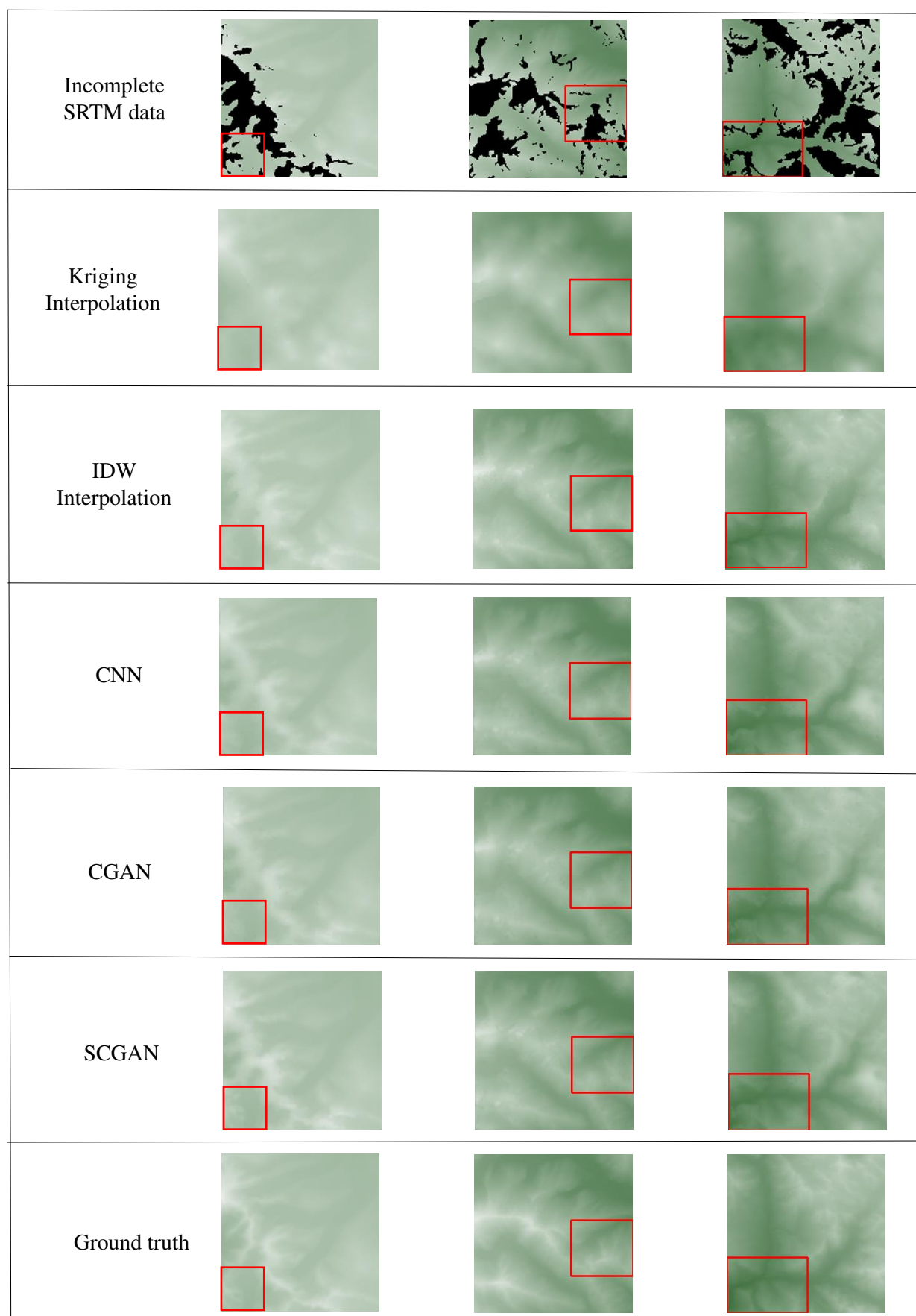


Fig. 7. SRTM 3 Arc-Second global data restoration results. Red boxes highlight regions in which the relative performance is most apparent.

8.1.1. Qualitative evaluations

Fig. 7 illustrates the void filling results obtained from the straightforward interpolation, the CNN, the baseline CGAN and the SCGAN. The first row in Fig. 7 displays the incomplete SRTM data. The second, third, fourth and fifth rows illustrate the void filling results obtained from different methods, separately. The bottom row shows the ground truth SRTM data. Each red box surrounds a local patch for which different methods represent different restoration performances. It is clear that the results from the SCGAN model agree best with the ground truth SRTM data among the four methods.

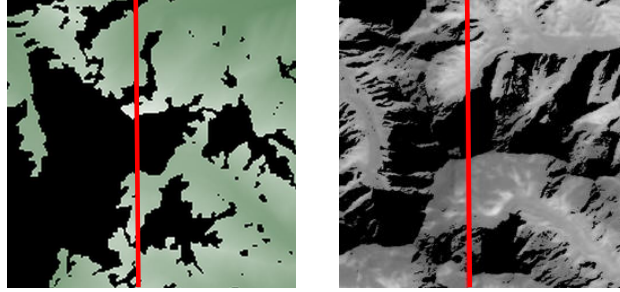


Fig. 8. The sectional view location marked in terms of a red line. The patch on the left is an SRTM 3 Arc-Second global data patch with void regions, and the patch on the right is its corresponding shadow map.

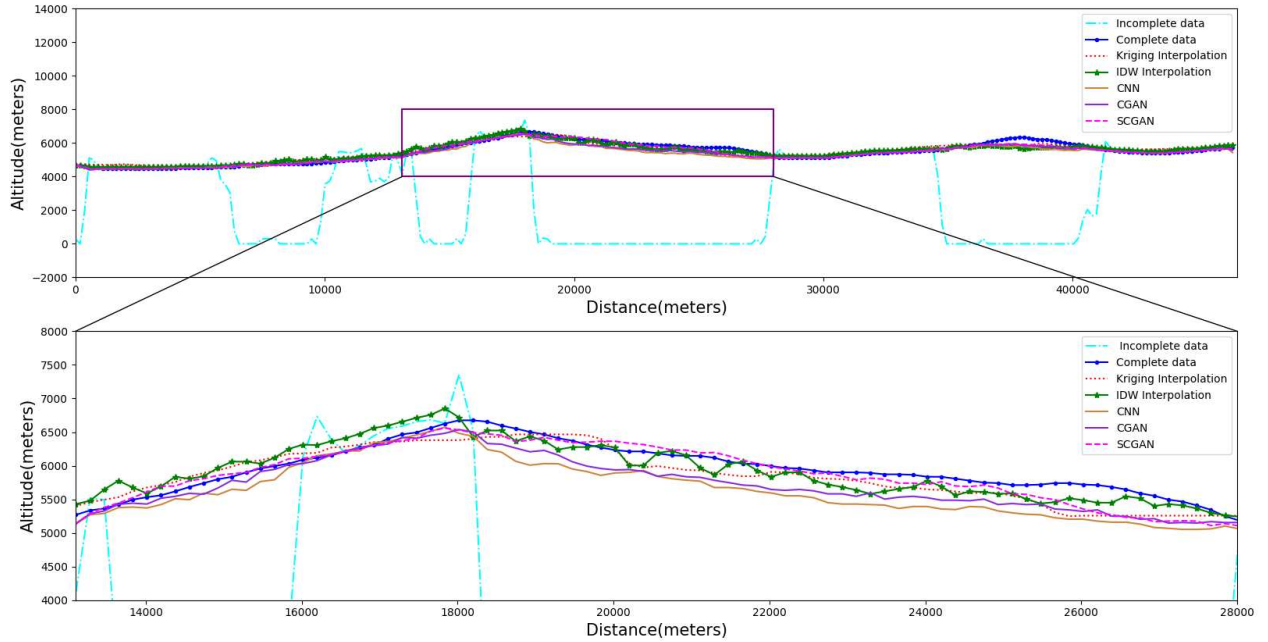


Fig. 9. The cross section view (for restored data along the red marked line in Fig. 8).

In order to further illustrate the performance of different methods qualitatively, the sectional view of a restored mountain curve along $85^{\circ}08'E$ at around $28^{\circ}19'N$ is investigated. The specific sectional view location is illustrated as the red line in Fig. 8. The sectional views of the mountain curve restored from

different methods are shown in Fig. 9. The top row shows the full sectional view and the bottom row displays the zoomed in restored curves with respect to one no-data shadow area. The altitude of the investigated region in Fig. 9 varies roughly from 5000 m to 6500 m. The Kriging interpolation tends to produce an over-smooth curve at the summit area but exhibits unfavorable fluctuations at relatively flat regions. The IDW interpolation produces an unfavorable abrupt change at the summit area. In contrast, the results obtained from all the deep learning based approaches vary more smoothly at the same area and show better resemblance to the complete data. It is clear that the SCGAN outperforms the interpolation methods, the CNN, and the baseline CGAN methods in terms of both the overall mountain curves and the shadowed mountain curves.

More specifically, in the bottom subfigure of Fig. 9, there are spikes on the top of the incomplete data curve. The spikes are considered as outliers and are supposed to be manually removed when applying most existing interpolation strategies. Our SCGAN method (as well as the alternative deep learning based methods) does not have such a requirement. It operates automatically without manual manipulations. The SCGAN restored curve varies smoothly in the neighborhood where spikes exist in the incomplete data. One reason for this advantage is that the deep learning strategy automatically learns the data variation heuristics that are against the appearance of outliers.

8.1.2. Quantitative evaluations

Table 1: PSNR (dB) for the overall investigated region (SRTM 3 Arc-Second global data).

Locations	Kriging	IDW	CNN	CGAN	SCGAN
$28^{\circ}31'N$ $85^{\circ}34'E$	38.14	38.48	42.41	42.47	43.93
$28^{\circ}23'N$ $85^{\circ}09'E$	37.96	37.18	37.19	37.44	40.14
$28^{\circ}21'N$ $85^{\circ}29'E$	36.86	37.53	37.88	37.97	38.89

Both peak signal to noise ratio (PSNR) and root mean square error (RMSE) are used for quantitatively evaluating the restoration accuracy. A larger PSNR value reflects better accuracy. In contrast, a smaller RMSE value reflects better accuracy. Table 1 shows the PSNR values of the two representative interpolation methods, the CNN, the baseline CGAN and the SCGAN, separately.

Table 2: RMSE (m) for the overall investigated region (SRTM 3 Arc-Second global data).

Locations	Kriging	IDW	CNN	CGAN	SCGAN
$28^{\circ}31'N$ $85^{\circ}34'E$	138.25	133.75	84.26	80.56	71.37
$28^{\circ}23'N$ $85^{\circ}09'E$	141.85	155.31	151.88	149.92	107.75
$28^{\circ}21'N$ $85^{\circ}29'E$	141.50	149.33	144.89	141.71	127.59

Table 3: PSNR (dB) for local shadowed regions (SRTM 3 Arc-Second global data).

Locations	Kriging	IDW	CNN	CGAN	SCGAN
$28^{\circ}31'N$ $85^{\circ}34'E$	38.99	39.42	42.23	42.24	43.13
$28^{\circ}23'N$ $85^{\circ}09'E$	37.59	38.79	41.82	42.28	42.90
$28^{\circ}21'N$ $85^{\circ}29'E$	36.86	37.28	38.43	38.85	43.82

Table 4: RMSE (m) for local shadowed regions (SRTM 3 Arc-Second global data).

Locations	Kriging	IDW	CNN	CGAN	SCGAN
$28^{\circ}31'N$ $85^{\circ}34'E$	125.37	125.33	86.57	86.57	53.42
$28^{\circ}23'N$ $85^{\circ}09'E$	148.31	131.83	95.73	86.34	53.71
$28^{\circ}21'N$ $85^{\circ}29'E$	167.53	152.33	136.69	128.45	103.38

Observations from Table 1 indicate that the CNN, the baseline CGAN and the SCGAN significantly outperform the interpolation methods. The key effective factor is that the deep learning methods characterize and learn the varying heuristics of mountains from the training data, and in contrast the interpolation methods do not explore the training data but just employ test data for restoration. Additionally, benefiting from incorporating the geometric shadow constraints into training the model, the SCGAN outperforms the CNN and the baseline CGAN.

Table 2 shows the RMSE values of different methods. Similar to those in Table 1, the CNN, the baseline CGAN and the SCGAN exhibit much better RMSE than the interpolation methods, and our SCGAN obtains the best RMSE among the four methods.

Table 1 and Table 2 evaluate the overall restoration results obtained from different methods. The SCGAN model achieves the best performance. To further evaluate these methods in detail, the PSNR and RMSE in shadow regions are calculated and compared among these different methods.

Table 3 represents the PSNR values in shadow regions obtained by different methods. The deep learning based methods (i.e., the CNN, the CGAN and the SCGAN) outperform the interpolation methods with higher PSNR. The CGAN model outperforms the CNN model due to the adversarially training process between the generator and the discriminator. Additionally, the shadow constraints are the key factors that enable our SCGAN model to achieve the best performance.

Table 4 shows the RMSE values in shadow regions obtained by different methods. Similar to Table 3, the SCGAN model outperforms the alternative comparison models (i.e., the interpolation methods, the CNN and the baseline CGAN).

The altitude of our study area is higher than 4,000 m, and even around 8,000 m in some regions. It

is observed that the valid SRTM 3 Arc-Second global data for the investigated area may suffer from errors larger than 100 m (Zhang et al., 2018). We consider an error around 100 m as an acceptable uncertainty for our study area because it is a small value relative to the investigated altitude as high as from 4,000 m to 8,000 m. Experimental results validate that our method effectively fills the SRTM data such that the restored data and the original valid data exhibit the same level of data quality in terms of error. Moreover, our method outperforms the comparison state of the art methods in terms of data restoration accuracy.

8.2. Evaluations based on SRTM 1 Arc-Second global data

In order to further validate the effectiveness of our SCGAN, we experimentally evaluate it with the 30 m resolution SRTM global data, i.e., the SRTM 1 Arc-Second global data². The SRTM 1 Arc-Second global data is almost complete with little voids. Applying the SCGAN model to refine the data with little voids cannot thoroughly evaluate its effectiveness. In order to extensively evaluate the capability of the SCGAN in data restoration, we manually delete local data which covers a fairly large regular region from the original data, artificially making the data considerably incomplete for testing the SCGAN model. Specifically, we delete the data surrounded by a box in the center of an investigated patch, and the deleted data covers 30% area of the patch. Different from the randomly shaped void regions in Section 8.1, the manually created void regions exhibit forms of a regular shape. The clear boundary between void and data regions makes it easier for us to examine the variation between filled data and existing data.

Different from the randomly located and arbitrarily shaped voids in the SRTM 3 Arc-Second global data used in Section 8.1, the voids in the SRTM 1 Arc-Second global data used here are artificially made with fixed locations and regular shapes. Moreover, the resolution of the SRTM 1 Arc-Second global data is higher than that of the SRTM 3 Arc-Second global data. In this scenario, we use the nearest neighbor (NN) interpolation method, which is widely used for restoring high-resolution data with regular size, as the representative interpolation method for experimental comparison.

8.2.1. Qualitative evaluations

Fig. 10 illustrates the void filling results obtained from different methods. The three learning based models, i.e., the CNN, the baseline CGAN, and the SCGAN outperform the NN interpolation, with the best results given by the proposed SCGAN model. More specifically, it is observed that the interpolation method is brute force such that the variation between the filled data and the existing data is visually very abrupt. In contrast, the SCGAN provides the visually smoothest variation between the filled data and the existing data, among the comparison methods.

In order to qualitatively evaluate the performance of different methods one step further, we investigate the

²<https://earthexplorer.usgs.gov/>

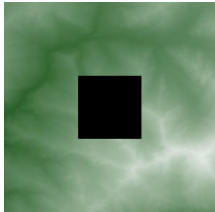
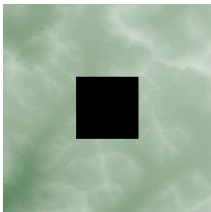
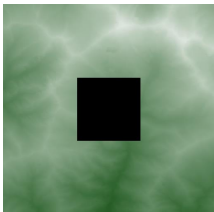
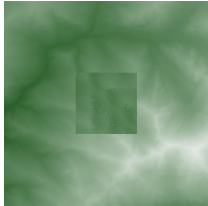

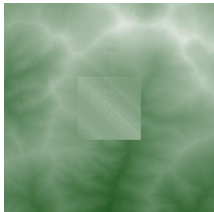
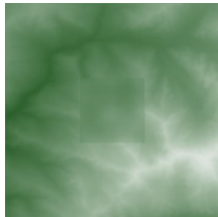
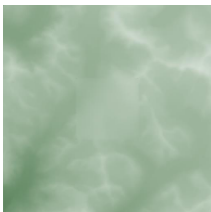
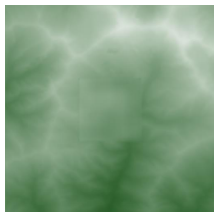
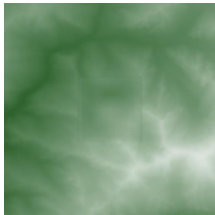
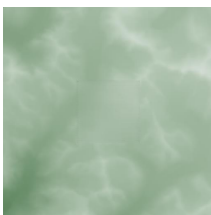
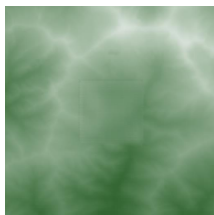
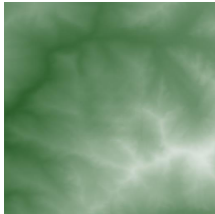
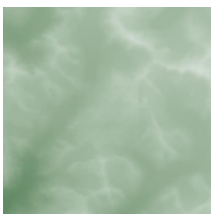
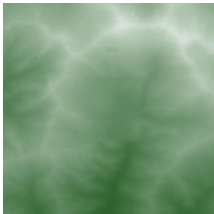
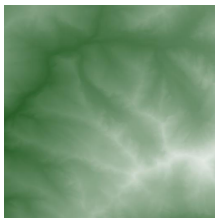

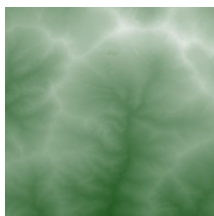
Incomplete SRTM data			
NN			
CNN			
CGAN			
SCGAN			
Ground truth			

Fig. 10. SRTM 1 Arc-Second global data restoration results.

sectional view of restored mountain curves along $85^{\circ}03'E$ at around $28^{\circ}13'N$. The location of the sectional view is marked as the red line in Fig. 11.

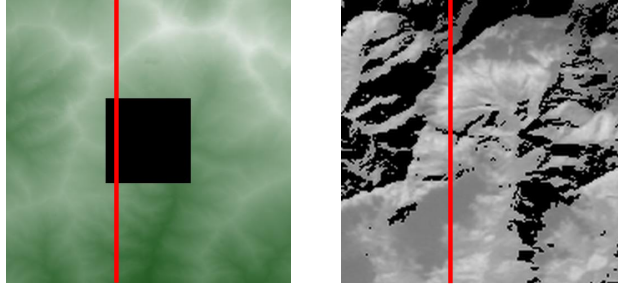


Fig. 11. The sectional view location marked in terms of a red line. The patch on the left is an SRTM 1 Arc-Second global data patch with manual data deletion, and the patch on the right is its corresponding shadow map.

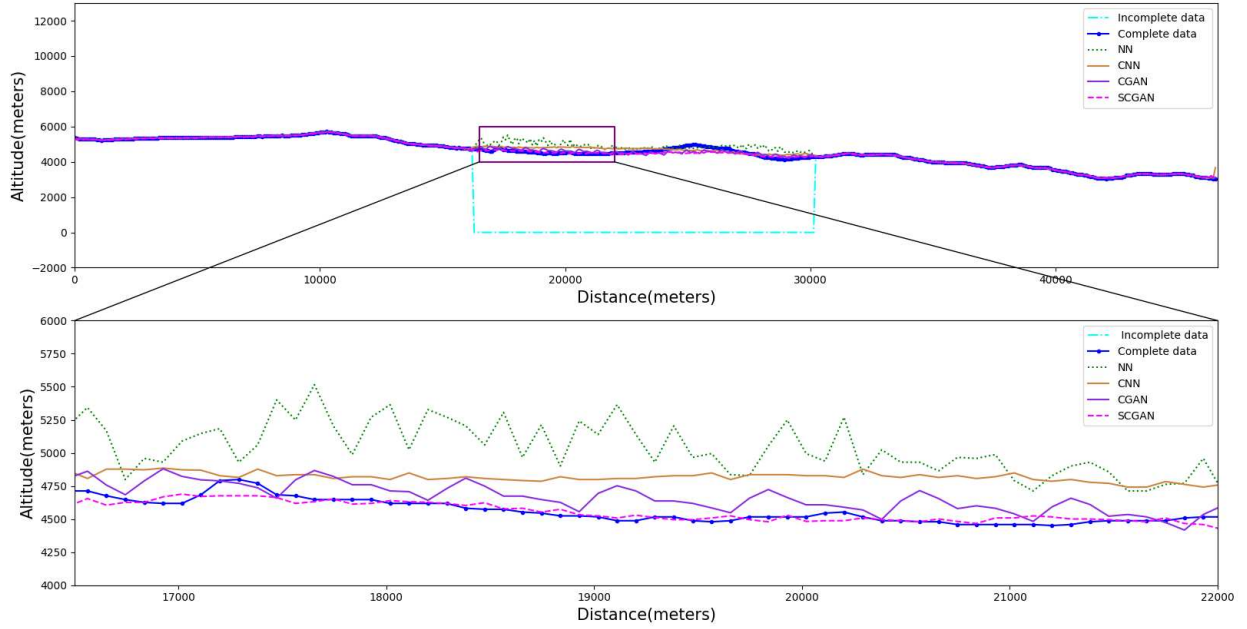


Fig. 12. The cross section view (for restored data along the red marked line in Fig. 11).

Fig. 12 illustrates the cross section view for the incomplete data, ground truth (complete data) and four restoration results. In Fig. 12, the top row shows the full sectional view and the bottom row displays the zoomed in restored curves within the void region. The NN interpolation method restores the artificial void region in the form of an inaccurate zigzag curve. This inaccuracy normally arises in large void regions where even the nearest neighboring available data samples are little related to the restored data samples. On the other hand, the curves restored based on the deep learning based methods suffer less from this shortcoming, because the deep learning methods have learned terrain variations from sufficient training data. Overall, it is clear that the SCGAN outperforms the interpolation, the CNN and the baseline CGAN methods in terms

of both the overall and the local zoomed in mountain curves.

8.2.2. Quantitative evaluations

Table 5: PSNR (dB) for the overall investigated region (SRTM 1 Arc-Second global data).

Locations	NN	CNN	CGAN	SCGAN
$28^{\circ}25'N$ $85^{\circ}23'E$	42.27	45.43	46.04	46.57
$28^{\circ}18'N$ $85^{\circ}27'E$	32.39	38.00	39.72	41.30
$28^{\circ}14'N$ $85^{\circ}05'E$	34.39	40.36	43.00	44.34

Table 6: RMSE (m) for the overall investigated region (SRTM 1 Arc-Second global data).

Locations	NN	CNN	CGAN	SCGAN
$28^{\circ}25'N$ $85^{\circ}23'E$	86.73	58.45	53.30	52.62
$28^{\circ}18'N$ $85^{\circ}27'E$	270.35	146.85	114.59	90.10
$28^{\circ}14'N$ $85^{\circ}05'E$	214.68	106.52	77.43	65.63

Tables 5 and 6 give the PSNR and RMSE evaluations of the overall restoration results, respectively. The three learning based models, i.e., the CNN, the baseline CGAN method and the proposed SCGAN model
495 outperform the interpolation method, with the best the best evaluation metric values given by the SCGAN model.

Table 7: PSNR (dB) for local shadowed regions (SRTM 1 Arc-Second global data).

Locations	NN	CNN	CGAN	SCGAN
$28^{\circ}25'N$ $85^{\circ}23'E$	47.45	49.92	49.12	50.24
$28^{\circ}18'N$ $85^{\circ}27'E$	41.45	42.15	42.79	43.67
$28^{\circ}14'N$ $85^{\circ}05'E$	52.73	51.96	52.11	52.28

Tables 7 and 8 give the PSNR and RMSE evaluations of the restoration results for local shadow regions, respectively. Though the three learning based models, i.e., the CNN, the baseline CGAN method and the proposed SCGAN model generally outperform the interpolation method, one interesting observation arises
500 that the NN interpolation achieves slightly better performance than the SCGAN in terms of both PSNR and RMSE for the patch at the location $28^{\circ}14'N85^{\circ}05'E$. Here the investigated patch and its corresponding shadow map are illustrated in Fig. 11. One possible reason for the result is that the shadows in the void data region just cover very small areas, and the interpolation method tends to be superior to the SCGAN

(and the other learning based methods) in terms of restoring small size data with limited available shadow information. In contrast, when tackling large region restoration (with both large void regions and large shadow regions), the SCGAN overwhelmingly outperforms the NN interpolation method, as observed in the experimental evaluations in Sections 8.1 and 8.2.

Table 8: RMSE (m) for local shadowed regions (SRTM 1 Arc-Second global data).

Locations	NN	CNN	CGAN	SCGAN
28°25′N 85°23′E	39.15	23.59	24.37	21.05
28°18′N 85°27′E	97.08	83.57	68.32	57.47
28°14′N 85°05′E	12.36	20.61	14.65	14.50

9. Conclusions

We have presented a shadow constrained generative adversarial neural network (SCGAN) for filling the voids of SRTM data in mountainous areas and thus obtained the restored SRTM data. Compared with straightforward deep learning models such as the convolutional neural network (CNN) and the conditional generative adversarial network (CGAN), the proposed SCGAN model has encoded geometric shadow constraints. Unlike the data driven strategy conducted via the straightforward CNN and CGAN, the geometric shadow constraints have endowed our SCGAN with knowledge of shadow cues. The geometric shadow constraints incorporated into the SCGAN have been in favor of restoring SRTM data satisfying the shadow cues. Therefore, the SCGAN has effectively addressed the incorrect restoration which might violate the geological shadowing rules. Empirical comparison results have validated that the SCGAN outperforms the interpolation, the CNN and the CGAN based methods. Additionally, training our proposed SCGAN model has integrated the space and in-situ measurements data. Therefore, the SCGAN has not only provided a new deep learning model that incorporates shadow geometry constraints but also formed a novel data fusion framework that integrates data from satellite, shuttle based and in-situ observations. The SCGAN is a general methodology for remote sensing surface restoration that takes advantage of the virtues of both shadow geometric cues and deep learning representations. Its application is not confined to filling SRTM voids but can be generalized to refining other kinds of remote sensing data. In the future, it is worth investigating how to extend our SCGAN model to other data refinement scenarios such as LiDAR DEM data restoration.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Project 61971444, Shandong Provincial Natural Science Foundation under Project ZR2019MF019, and Fundamental Research Funds for the Central Universities under Project 18CX05014A. William A. P. Smith is a recipient of the RAEng/Leverhulme Trust Senior Research Fellowship.

References

- Antipov, G., Baccouche, M., & Dugelay, J. (2017). Face aging with conditional generative adversarial networks. In *IEEE International Conference on Image Processing* (pp. 2089–2093).
- Arun, P. (2013). A comparative analysis of different dem interpolation methods. *The Egyptian Journal of Remote Sensing and Space Science*, 16, 133 – 139.
- Boncori, J. P. M. (2016). Caveats concerning the use of srtm dem version 4.1 (cgia-rcsi). *Remote Sensing*, 8, 793.
- Chan, T. F., Sandberg, B., & Vese, L. A. (2000). Active contours without edges for vector-valued images. *Journal of Visual Communication and Image Representation*, 11, 130 – 141.
- Chandraker, M., Agarwal, S., & Kriegman, D. (2007). Shadowcuts: Photometric stereo with shadows. In *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–8).
- Chung, K., Lin, Y., & Huang, Y. (2009). Efficient shadow detection of color aerial images based on successive thresholding scheme. *IEEE Transactions on Geoscience and Remote Sensing*, 47, 671–682.
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 35, 53–65.
- Dong, G., Chen, F., & Ren, P. (2018). Filling srtm void data via conditional adversarial networks. In *IEEE International Geoscience and Remote Sensing Symposium* (pp. 7441–7443).
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45, 361.
- Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2414–2423).
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* (pp. 2672–2680).
- Grohman, G., Kroenung, G., & Strebeck, J. (2006). Filling srtm voids: The delta surface fill method. *Photogrammetric Engineering and Remote Sensing*, 72, 213–216.
- Hall, O., Falorni, G., & Bras, R. L. (2005). Characterization and quantification of data voids in the shuttle radar topography mission data. *IEEE Geoscience and Remote Sensing Letters*, 2, 177–181.

- Heritage, G. L., Milan, D. J., Large, A. R., & Fuller, I. C. (2009). Influence of survey strategy and interpolation model on dem quality. *Geomorphology*, *112*, 334–344.
- Hirt, C. (2018). Artefact detection in global digital elevation models (dems): The maximum slope approach and its application for complete screening of the srtm v4.1 and merit dems. *Remote Sensing of Environment*, *207*, 27 – 41.
- Hogan, J., & Smith, W. A. P. (2010). Refinement of digital elevation models from shadowing cues. In *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1181–1188).
- Isola, P., Zhu, J., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5967–5976).
- Jafarzadegan, K., & Merwade, V. (2017). A dem-based approach for large-scale floodplain mapping in ungauged watersheds. *Journal of hydrology*, *550*, 650–662.
- Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). Hole-filled SRTM for the globe Version 4. In *Consortium for Spatial Information*.
- Karkee, M., Steward, B. L., & Aziz, S. A. (2008). Improving quality of public domain digital elevation models through data fusion. *Biosystems Engineering*, *101*, 293 – 305.
- Kellndorfer, J., Walker, W., Pierce, L., Dobson, C., Fites, J. A., Hunsaker, C., Vona, J., & Clutter, M. (2004). Vegetation height estimation from shuttle radar topography mission and national elevation datasets. *Remote Sensing of Environment*, *93*, 339 – 358.
- Lee, C., Oh, J., Hong, C., & Youn, J. (2015). Automated generation of a digital elevation model over steep terrain in antarctica from high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *53*, 1186–1194.
- Li, C., Kao, C.-Y., Gore, J. C., & Ding, Z. (2008). Minimization of region-scalable fitting energy for image segmentation. *Trans. Img. Proc.*, *17*, 1940–1949.
- Liao, J., Buchholz, B., Thiery, J., Bauszat, P., & Eisemann, E. (2017). Indoor scene reconstruction using near-light photometric stereo. *IEEE Transactions on Image Processing*, *26*, 1089–1101.
- Ling, F., Zhang, Q., & Wang, C. (2007). Filling voids of srtm with landsat sensor imagery in rugged terrain. *International Journal of Remote Sensing*, *28*, 465–471.
- Lu, C., & Drew, M. S. (2005). Shadow segmentation and shadow-free chromaticity via markov random fields. In *Color and Imaging Conference* (pp. 125–129).

Milan, D. J., Heritage, G. L., Large, A. R., & Fuller, I. C. (2011). Filtering spatial error from dems:

Implications for morphological change estimation. *Geomorphology*, 125, 160 – 171.

Mukul, M., Srivastava, V., Jade, S., & Mukul, M. (2017). Uncertainties in the shuttle radar topography mission (srtm) heights: Insights from the indian himalaya and peninsula. *Scientific Reports*, 7, 41672.

Polidorio, A. M., Flores, F. C., Imai, N. N., Tommaselli, A. M. G., & Franco, C. (2003). Automatic shadow segmentation in aerial color images. In *Brazilian Symposium on Computer Graphics and Image Processing* (pp. 270–277).

Reuter, H. I., Nelson, A., & Jarvis, A. (2007). An evaluation of void-filling interpolation methods for srtm data. *International Journal of Geographical Information Science*, 21, 983–1008.

Rizzoli, P., Martone, M., Gonzalez, C., Wecklich, C., Tridon, D. B., Bräutigam, B., Bachmann, M., Schulze, D., Fritz, T., Huber, M. et al. (2017). Generation and performance assessment of the global tandem-x digital elevation model. *Journal of Photogrammetry and Remote Sensing*, 132, 119–139.

Rodriguez, E., Morris, C. S., & Belz, J. E. (2006). A global assessment of the srtm performance. *Photogrammetric Engineering and Remote Sensing*, 3, 249 – 260.

Rodriguez, E., Morris, C. S., Belz, J. E., Chapin, E., Martin, J., Daffer, W., & Hensley, S. (2005). An assessment of the srtm topographic products. *Technical Report JPLD-31639, Jet Propulsion Laboratory*, 143.

Shelhamer, E., Long, J., & Darrell, T. (2017). Fully convolutional networks for semantic segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39, 640–651.

Teke, M., Başeski, E., Ok, A. Ö., Yüksel, B., & Şenaras, Ç. (2011). Multi-spectral false color shadow detection. In *Photogrammetric Image Analysis* (pp. 109–119).

Toutin, T. (2008). ASTER DEMs for geomatic and geoscientific applications: a review. *International Journal of Remote Sensing*, 29, 1855–1875.

Villarini, B., Gkelias, A., & Argyriou, V. (2017). Photometric stereo for 3d face reconstruction using non linear illumination models. *Lecture Notes in Computer Science*, (pp. 140–152).

Wendleder, A., Felbier, A., Wessel, B., Huber, M., & Roth, A. (2016). A method to estimate long-wave height errors of srtm c-band dem. *IEEE Geoscience and Remote Sensing Letters*, 13, 696–700.

Wu, C., Narasimhan, S. G., & Jaramaz, B. (2010). A multi-image shape-from-shading framework for near-lighting perspective endoscopes. *International Journal of Computer Vision*, 86, 211–228.

Yang, K., Smith, L. C., Chu, V. W., Gleason, C. J., & Li, M. (2015). A caution on the use of surface digital elevation models to simulate supraglacial hydrology of the greenland ice sheet. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 5212–5224.

Yu, X., Zhang, H., Luo, C., Qi, H., & Ren, P. (2018). Oil spill segmentation via adversarial f -divergence learning. *IEEE Transactions on Geoscience and Remote Sensing*, 56, 4973–4988.

Yue, L., Shen, H., Zhang, L., Zheng, X., Zhang, F., & Yuan, Q. (2017). High-quality seamless dem generation blending srtm-1, aster gdem v2 and icesat/glas observations. *Journal of Photogrammetry and Remote Sensing*, 123, 20–34.

Yue, L., Yu, W., Shen, H., Zhang, L., & He, Y. (2015). Accuracy assessment of srtm v4.1 and aster gdem v2 in high-altitude mountainous areas: A case study in yulong snow mountain, china. In *IEEE International Geoscience and Remote Sensing Symposium* (pp. 5011–5014).

Zhan, Y., Hu, D., Wang, Y., & Yu, X. (2018). Semisupervised hyperspectral image classification based on generative adversarial networks. *IEEE Geoscience and Remote Sensing Letters*, 15, 212–216.

Zhang, Q., Yang, Q., Cheng, J., & Wang, C. (2018). Characteristics of 3 srtm errors in china. *Geomatics and Information Science of Wuhan University*, 43, 684–690.

Zhu, A., Meng, Y., & Zhang, C. (2017). An improved adam algorithm using look-ahead. In *International Conference on Deep Learning Technologies* (pp. 19–22).

Table 9: List of Figure Captions

Figure 1	A sample of data.
Figure 2	SRTM data restoration by the CGAN and SCGAN models.
Figure 3	Limitations of the CGAN model. The incorrect restoration would be apparent if cast shadows were not or inappropriately considered.
Figure 4	Mountain shadow geometry.
Figure 5	Shadow segmentation.
Figure 6	SRTM data restoration via the generator of the SCGAN.
Figure 7	SRTM 3 Arc-Second global data restoration results. Red boxes highlight regions in which the relative performance is most apparent.
Figure 8	The sectional view location marked in terms of a red line. The patch on the left is an SRTM 3 Arc-Second global data patch with void regions, and the patch on the right is its corresponding shadow map.
Figure 9	The cross section view (for restored data along the red marked line in Fig. 8).
Figure 10	SRTM 1 Arc-Second global data restoration results.
Figure 11	The sectional view location marked in terms of a red line. The patch on the left is an SRTM 1 Arc-Second global data patch with manual data deletion, and the patch on the right is its corresponding shadow map.
Figure 12	The cross section view (for restored data along the red marked line in Fig. 11).